County-Level Estimates of the Wage Prospects
of Low-Skill Workers

David C. Ribar[*]
Department of Economics
The George Washington University
2201 G Street, NW
Washington, District of Columbia 20052
E-mail: dcr7@gwu.edu

December 1999
Preliminary, please do not cite without permission

County-Level Estimates of the Wage Prospects of Low-Skill Workers

1. Introduction

It is difficult to characterize the performance of the U. S. economy over the last decade. By some measures, the economy has done extremely well. With the exception of a brief recession in 1990-91, the economy as a whole has experienced steady growth, rising productivity, low and falling unemployment, and little inflation. Following sharp run-ups at the start of the decade, there have also been substantial declines in welfare caseloads and poverty. By several other measures, however, the economy's performance has been disappointing. Inflation-adjusted wages and earnings are essentially unchanged from the start of the decade, and poverty rates, while recently falling, remain higher than levels recorded during the 1970s.

The economy has also generated very different outcomes for people in different parts of the country and with different levels of market skills. Average annual growth from 1990-94 was three times higher in the western states than in the northeastern states, and poverty rates in some states have been twice as high as the rates in other states. Wages and earnings for workers with low levels of schooling, which generally declined through the 1980s, continued to fall until about 1995 and only recently began to recover. Figure 1 shows the trends in average annual earnings from 1989-98 for all workers and for workers without college.

Characterizing the labor market outcomes for low-skill workers in different areas of the country has become increasingly important in the aftermath of welfare reform. The Personal Responsibility and Work Opportunity Reconciliation Act (PRWORA) of 1996 was enacted with the goals of decreasing

individuals' dependence on government assistance programs, promoting actions that lead to economic self-sufficiency, and increasing the flexibility of states in providing cash assistance, training, and employment support to the poor. The legislation emphasized employment as a route out of poverty by incorporating stringent work and work-readiness requirements for the beneficiaries of different programs and backing these requirements up with strict time limits and other benefits sanctions. With large numbers of recipients now reaching these time limits, it is important to know whether (and which) local labor markets can accommodate the influx of low-skill workers and provide wages that will lead to self-sufficiency.

This paper examines low-skill wage opportunities for women at the county level over the period 1989-96. Currently, reliable direct measures of wage measures for different demographic and skill groups are only available for large geographic areas such as regions and populous states or at ten-year intervals (from the Decennial Census) for some smaller areas. This study constructs indirect annual measures for all counties from 1989-96 by combining skill-specific information on earnings and employment from the Sample Edited Detail File (SEDF) of the 1990 Decennial Census and the 1991-97 Annual Demographic files of the Current Population Survey (CPS) with annual industry-specific information from the Regional Economic Information System (REIS). Special versions of the SEDF and CPS files, which identify county of residence and work, are used.

The study regresses the low-skill wage data for women from the SEDF and CPS files on a set of personal variables from the combined files and local employment and earnings measures derived the REIS. The wage regressions are corrected for selectivity from the employment decision and account for county-specific effects as well as general time effects. Estimates from the regressions are then

combined with the available employment and earnings data from the REIS to impute wages for low-skill women workers across counties.

The remainder of this paper is organized as follows. Section 2 briefly reviews previous research that has examined the effects of local labor market conditions on earnings for low-skill workers. Section 3 describes the individual- and local-level data used in the empirical analysis. Regression estimates are reported in Section 4, and wage imputations derived from the regressions are examined in Section 5. Section 6 discusses the findings and concludes.

2. Previous research

Numerous studies have examined the effects of local labor market conditions on wages, earnings and related outcomes. For brevity, my review focuses only on the subset of studies that have considered outcomes for low-skill workers.

Much of the literature has considered specific groups of low-skill workers. For instance, Freeman (1981) and Freeman and Rodgers (1999) examined earnings among young men with low levels of schooling and found that their earnings were adversely affected by high general rates of metropolitan unemployment. Bartik and Eberts (1999), Card and Lemieux (1997), Council of Economic Advisors (1997, 1999), Figlio and Ziliak (1999), and Fitzgerald (1995) examined the effects of labor market conditions on welfare receipt. With the exception of the study by Card and Lemieux, these studies found that tight labor markets reduced recipiency. Related studies by Bartik (1993) and Walters (1990) reported that high levels of labor demand reduced poverty.

The remaining studies have compared the effects of local labor conditions on the earnings of

different types of workers. Topel (1986) found that employment growth and positive current employment shocks increased men's wages and that the effects were stronger for men with low levels of education. Hoynes (1999) similarly reported that low-skill workers' earnings, and especially low-skill women's earnings, responded more to labor market shocks than high-skill workers. Bartik (1999) found that the earnings of men who were not college graduates and women who were neither college graduates or single heads of households were more sensitive to employment shocks than the earnings of more-educated men or women. However, he also found that the earnings of female non-graduates who were also single heads of households were less sensitive to shocks than those of other women.

Katz and Murphy (1992) found that shifts in demand were largely responsible for recent changes in the wage structure for skilled and unskilled workers. Juhn, Murphy and Pierce (1993) provided supporting evidence that increases in the returns to skills contributed to changes in the wage structure. In contrast, Bound and Holzer (1993) found that while industrial shifts (specifically, the decline in manufacturing employment) had a disproportionate negative effect on the employment of low-skill men, these shifts had little effect on their overall earnings.

Of the studies described above, only those by Bartik (1999), Hoynes (1999), and Katz and Murphy (1992) considered low-skill women's earnings. Almost all of the studies have considered labor market outcomes for relatively large areas such as major metropolitan areas (Bartik 1993; Bound and Holzer 1993; Freeman 1982; Freeman and Rodgers 1999; Hoynes 1999), states (Bartik 1999; C.E.A. 1997, 1999; Figlio and Ziliak 1999; Topel 1986), regions (Card and Lemieux 1997), or the nation as a whole (Katz and Murphy 1992); only one (Fitzgerald 1995) has examined conditions at the county level. Many of the studies used aggregated data and thus included only limited controls for

4

demographic differences across workers. None of the studies accounted for potential selectivity in observed earnings from the employment decision.

3. Data

The primary data for this analysis come from the Sample Edited Detail File of the 1990 Decennial Census and confidential versions of the 1991-97 Annual Demographic (March) Supplements of the Current Population Survey (CPS). The information in the SEDF comes from the "long forms" which were administered as part of the 1990 Decennial Census. Thus, it represents a very large (one-in-six) cross-section sample of the U.S. population. The March files of the CPS are smaller and sample roughly 60,000 households per year. Detailed geographic information is attached to both the SEDF and confidential versions of the CPS. I obtained access to these files through a NSF/ASA/Census Bureau research fellowship.

Individual-level variables. The SEDF and March files of the CPS record comparable information on whether a person was employed during the previous year, the number of weeks the person worked, the number of hours worked in an average week, and the amount of money earned from different sources. From these measures, I construct three variables: a dummy variable indicating employment during the previous year, weekly earnings during the previous year (total personal earnings divided by weeks worked), and hourly earnings during the previous year (total personal earnings divided by weeks and typical hours worked). Nominal amounts were re-expressed in constant 1998 dollars using the Consumer Price Index for Urban Consumers (CPI-U).

In addition to the economic variables, the SEDF and CPS also contain comparable information

on the person's sex, age, ethnic origin, and schooling level. I use the sex and age information as

recorded. From the ethnic origin data, I construct dummy indicators for people of African origin and

people of other non-European origins (mostly native Americans, Asians, and Pacific Islanders); the

omitted category is European origin. I also construct a separate dummy variable for Hispanic origin

which may overlap with the other racial/ethnic categories. Using the schooling information, I distinguish

between four types of people: those who have not completed high school, those who have completed

high school (or equivalent) but have not gone on to college, those who have completed some college,

and those who have graduated college. Most of the empirical analysis focuses on women from the two

lowest education groups.

From the combined data set, I select non-institutionalized, civilians who were 16 to 64 years of

age. I then make a number of data exclusions. First, I exclude individuals below age 24 who were

enrolled in school. Second, I exclude observations if the person's hours or earnings were top-coded or

if the calculation of weekly or hourly earnings was unreasonable – below $1 or above $250 for hourly

wages and below $1 or above $10,000 for weekly wages. Third, I drop observations with allocated

economic or demographic data.[1] Even with these exclusions, the resulting data set still contains several

million observations. To make the data more manageable, I randomly sample observations from the

SEDF for counties with more than 100,000 residents (the sampling probability is 100,000/population)

and re-weight the remaining observations accordingly. All of the statistical analyses incorporate

sampling weights scaled to the annual sample sizes.

Means and standard deviations for the individual-level variables drawn from the SEDF and CPS

files are reported in Appendix A1. The appendix lists statistics separately for subsamples of women

6

who did not complete high school (394,701 observations) and women who completed high school but did not go on to college (1,185,657 observations).

Measures of local labor market conditions. Using the geographic identifiers in the SEDF and CPS files, I can link the individual-level observations to measures of local labor market conditions. An important issue that must be addressed first, though, is the definition of the market itself – what are the geographic boundaries and what types of labor are involved?

There is little agreement in the existing empirical literature regarding what constitutes a "local" labor market. If we base our definition on the measures that have actually been employed in research, a fair definition of "local" would be any geographic area smaller than a Census region. As mentioned earlier, almost all of the previous studies have used states or metropolitan areas. While research on commuting patterns by urban and regional economists suggests that a narrower geographic definition should be adopted, labor market studies have generally not considered smaller areas or incorporated information on commuting.[2]

Some data for smaller areas are available. For instance, county-level unemployment estimates are reported in the Geographic Profile of Employment and Unemployment from the Bureau of Labor Statistics, and industry-level earnings and employment estimates are reported in the Regional Economic Information System (REIS) from the Bureau of Economic Analysis. However, these data have serious limitations. In the case of the local unemployment figures, the data are very broad and, thus, may not reflect opportunities for specific skill or occupation groups. In the case of the REIS employment and earnings measures, the data are specific to industries rather than skill or occupation groups and are recorded on a place of work rather than a place of residence basis.

Data on skill-specific opportunities can be estimated using data from the CPS and the SEDF. Both files contain information on worker characteristics including age and education that can be used to infer skill levels as well as data on labor market outcomes. However, once again there are limitations. The CPS can be used to construct annual or even monthly statistics for the nation as a whole and for some large subnational areas; however, there are not enough observations to generate reliable estimates for small areas. The SEDF does include enough observations to produce reliable direct measures of skill- and residence-specific labor market outcomes for small areas but is only available at ten-year intervals.

As an initial analysis to inform the construction of the local labor market measures, I use the SEDF to examine where and in which industries employees with different skill levels work. Table 1 reports summary statistics from the SEDF on the percentages of women and men with different education levels that (a) work and live in the same county and (b) work in various industries. As the figures indicate, commuting and industry patterns vary substantially across gender and skill groups.

On average, men are more likely than women to travel beyond their county of residence to get to work. Commuting across counties also increases with an employee's level of schooling. Women with the least schooling have the lowest tendency to commute; nevertheless, a substantial number still do commute. Nearly one-sixth of women with less than a high school diploma and one-fifth of women with a diploma (or equivalent) but no college commute across counties.

Table 1 also indicates that the employment of women with low levels of schooling is concentrated in three industries: manufacturing, retail trade, and services. More than four-fifths of women with less than a high school diploma are employed in these three industries with nearly a third working in the low-

8

paid retail sector. Among women with a high school diploma but no college, two-thirds work in these three industries. While manufacturing, retail trade, and services account for more than half the employment among low-skilled men, a fourth industry – construction – is also important. The "big three" also account for just over half of the employment of college-educated men and women; however, the concentrations within specific industries differ substantially from those for less-educated workers, and other sectors such as public service account for a large share of employment.

The observed differences in industry employment and commuting patterns suggest that general population-wide aggregates and simple within-county measures will not accurately describe employment opportunities within gender and skill groups. Therefore, to form the study's labor market variables, I combine annual county-level place-of-work industry employment and earnings information from the REIS with employment weights derived from the SEDF to form county-of-residence measures of skill-specific opportunities. The weighting approach is similar to that developed by Bowen and Finegan (1969) and adopted in numerous subsequent studies to map industry employment data to demographic, skill or occupation groups; to my knowledge, though, the procedure has not been used in research studies to account for commuting patterns or define geographic labor markets.

In particular, I obtain total employment and earnings in one-digit S.I.C. industries for each county from 1989 through 1996. The industry-specific data in the REIS are based primarily on administrative records (ES-202 forms) submitted by employers to state employment agencies and are reported on a place of work rather than a place of residence basis.[3] Let $E_{\text{REIS}}(j, c, t)$ denote the total number of employees and $Y_{\text{REIS}}(j, c, t)$ denote the total earnings in industry j $(= 1, J)$ in county c $(= 1, C)$ in year t derived from the REIS.

To re-weight these data, let $e_{\text{SEDF}}(s, r, j, c)$ denote the fraction of employees in industry j and county c with skills s who commute from county r as estimated from the SEDF. I estimate annual skill- and residence-specific employment for year t using the weighting formula

$$\hat{M}(s,r,t) = \sum_{j=1}^{J} \left(\sum_{c=1}^{C} e_{\text{SEDF}}(s,r,j,c) E_{\text{REIS}}(j,c,t) \right). \tag{1}$$

In the empirical analysis, I re-express this figure in per capita terms by dividing through by the total population in the county of residence. I estimate annual skill- and residence-specific earnings per worker for year t using the formula

$$\hat{R}(s,r,t) = \hat{M}(s,r,t)^{-1} \sum_{j=1}^{J} \left(\sum_{c=1}^{C} e_{\text{SEDF}}(s,r,j,c) Y_{\text{REIS}}(j,c,t) \right). \tag{2}$$

Table 2 reports the national trends from 1989 to 1996 in earnings and employment using different measures. The first row of Table 2 lists total civilian employment per capita as reported in the REIS; the next three rows list per capita employment in manufacturing, retail trade, and services per capita also as reported in the REIS. The figures indicate the effects of the recession in the early 1990s with total employment declining in 1991 and 1992, beginning to rebound in 1993, and moving beyond its pre-recession peak by 1994. The figures also show the general decline in manufacturing employment and the overall growth in retail trade and service employment over the period.

The next four rows in Table 2 report average earnings per worker for all civilians and earnings per worker in manufacturing, retail trade, and services. As with the preceding figures, all of these numbers are derived directly from the REIS. The figures have also been adjusted for inflation using the CPI-U. The numbers indicate that total earnings per worker fell slightly during the recession but were

10

otherwise stagnant from 1989 to 1996. Across industries, earnings per worker in manufacturing rose over the period while earnings in retail trade fell.

The remaining rows show the weighted estimates of employment and earnings from equations (1) and (2) for women with different schooling levels. As expected, the weighted earnings figures increase with the level of schooling. However, the amount of increase is modest reflecting the fact that the estimates do not account for within-industry changes in skill-specific wages. The figures suggest that jobs for women who had not completed high school grew slightly (about 3 percent) from 1989 to 1990 while earnings were essentially flat. Jobs for women with higher levels of schooling appear to have increased more over the period (5 percent for high school graduates and 7 percent for those who went on to college). Earnings for more highly educated women also appear to have grown slightly (2 to 3 percent) over this period.

4. Regression analysis

As the preceding analysis shows, the employment-weighted labor market variables capture some of the variation in opportunities for specific skill groups. Although the variables are an improvement over general aggregates, they are still limited because they only indirectly incorporate the SEDF information and do not account for within-industry changes in employment opportunities.

To examine employment opportunities more directly, I use the combined individual-level data from the SEDF and CPS linked to the employment-weighted local labor market variables to estimate a series of skill- and region-specific log wage regressions of the form

$$\ln W(i,s,r,k,t) = \beta_1(s,k)\hat{M}(s,r,t) + \beta_2(s,k)\ln\hat{R}(s,r,t)$$
$$+ \beta_3(s,k)\big(\hat{M}(s,r,t) - \hat{M}(s,r,t-1)\big) \tag{3}$$
$$+ \beta_X(s,k)X(i,s,r,t) + \delta(s,k,t) + \mu(s,r) + \varepsilon(i,s,r,k,t).$$

In equation (3), k denotes the region of residence; $X(i, s, r, t)$ is a vector of individual-specific attributes; the

$\beta(s, k)$ terms are skill- and region-specific coefficients; $\delta(s, k, t)$ is a skill-, region-, and year-specific

effect; $\mu(s, r)$ is a skill- and county-of-residence specific effect; and $\varepsilon(i, s, r, k, t)$ is an individual-

specific error. The county-specific effects and the general time effects are estimated as fixed effects.

The coefficients on the individual-specific attributes are identified separately from the county and time

effects because the attributes vary within counties and over time. The county-specific economic

variables are identified because they vary over time for the subset of counties that are observed in the

March files of the CPS.

The wage specification was selected after some experimentation. Initial tests revealed that the

inclusion of county fixed effects significantly improved the fit of the model relative to models with state

effects or no effects. Initial tests also supported the inclusion of general time effects and the use of

separate regressions across regions. Although the data allow me to estimate separate regressions

below the regional level (e.g., state-specific regressions), the CPS files do not have enough

observations to provide reliable estimates once the data are stratified into skill- and gender-specific

components.

All of the specifications in the empirical analysis are also corrected for selectivity from the

employment decision using Heckman's (1979) two-stage technique. The first stage employment probit

models for the correction are specified analogously to equation (3). The inclusion of fixed effects in the

probit models leads to potentially inconsistent estimates because of the incidental parameters problem; however, because of the large numbers of observations for most of the counties, this should not be a serious issue. Initial specification tests supported the inclusion of the selectivity adjustments. Results for the first-stage employment probit models are reported in Appendix Tables A2 and A3.[4]

Coefficient estimates and heteroskedasticity-consistent standard errors for the region-specific regressions for women without a high school diploma are reported in Table 3. The first four columns of Table 3 list results from regressions that use the log of hourly earnings as the dependent variable; the next four columns list results from regressions the use the log of weekly earnings as the dependent variable. For brevity, the tables only report results for the economic variables, individual attributes, and selectivity term (8).

In the hourly earnings regressions, the estimated coefficient for the weighted county employment variable is positive for all four regions and significant for three of the regions. The magnitudes of the employment coefficients vary greatly across areas – the coefficients for the Northeast and South imply wage elasticities of about .3 while the coefficient for the Midwest implies a wage elasticity of 1.3. The coefficient on employment growth is negative for all four regions but also varies substantially from region to region. The coefficients for the county earnings variables are strongly negative and significant for the Northeast and Midwest and positive but insignificant in the South and West. The significant negative results for the earnings and employment growth variables are counter-intuitive. One possible explanation of these results is that increases in earnings and employment growth within local areas over time might reflect increases in the demand for other competing types of labor.

Among the other variables in the hourly earnings regressions, age has a significantly positive

13

coefficient, and age squared has a significantly negative coefficient in all four regions. The estimates for all four regions imply that hourly earnings for women without high school diplomas increase with age until about age 40 and decline thereafter. The coefficients for the racial and ethnic minority indicators are negative and generally significant across the regions. The point estimates imply particularly strong effects of minority status in some regions. For instance, African origin is estimated to reduce hourly earnings by 35 percent in the Midwest; Hispanic origin is estimated to reduce hourly earnings by a similar percentage in the Northeast.

The estimation results for the weekly earnings regressions in Table 3 are broadly similar to the results for the hourly earnings regressions; however, there is more variation in the coefficients across regions in the weekly earnings regressions and fewer significant estimates. The coefficient for the local employment variable is strongly positive and significant for the Midwest and West but insignificant in the Northeast and South. The area earnings variable has significant negative coefficients in the Northeast and Midwest, and the employment growth variable has a significant negative coefficient in the Midwest; these variables have insignificant coefficients elsewhere.

Table 4 reports similar hourly and weekly earnings regression results for women who have completed high school but not gone on to college. The local employment variable is estimated to have significant positive effects on both hourly and weekly earnings in all four regions. The implied elasticities range from about .5 in the Midwest to about 1.0 in the West for hourly earnings and from about .3 in the Northeast to about .9 in the West for weekly earnings. Local average earnings is estimated to have significantly negative effects on individual earnings in the Northeast and Midwest and insignificant negative effects elsewhere. Employment growth has significantly negative coefficients for hourly and

weekly earnings in the Northeast but mixed and insignificant coefficients elsewhere. Most of the coefficients for age, race, and ethnicity are significant with signs similar to the corresponding coefficients from Table 3. The age coefficients generally imply that the returns to experience are weaker for women who are high school graduates than for less educated women.

Alternative regressions incorporating direct county-level measures from the REIS of (a) total civilian employment, average annual earnings, and employment growth and (b) industry-specific employment, earnings, and employment growth for manufacturing, retail trade, and services have also been estimated. The alternative models produced nearly identical fits of the data and estimates that were qualitatively similar to the results from Tables 3 and 4. The reported results appear to be robust to respecification of the local labor market variables.

5. Comparison of wage imputations

To examine the general properties of the regression results, I use the estimates from Tables 3 and 4 to impute hourly earnings over all counties from 1989-96 for women who had and had not completed high school. To abstract from demographic differences across areas, I use the regression estimates to impute hourly earnings for a representative woman in each county who is white, non-Hispanic, and 30 years of age. The earnings imputations, weighted for county population size, are then averaged across counties to form regional aggregates.

Figure 2 shows how the earnings for women without high school diplomas have changed since 1989 in each of the four regions according to three different measures. The thick solid lines in the graphs are the average imputed hourly earnings from the Table 3. The thin solid lines are average

imputed hourly earnings based on regressions (not reported) that omit corrections for employment

selectivity. The thin dashed lines represent the employment-weighted annual earnings. For

comparability, all of the earnings figures are expressed as a percentage of their 1989 values.

The graphs from Figure 2 indicate that the regression-based imputations of hourly wages for each

region are slightly, but not overly, sensitive to corrections for employment selectivity. While there are

some differences, the imputations based on the corrected and uncorrected estimates are very close in

most years.

The regression-based imputations do, however, differ markedly from the employment-weighted

earnings estimates. All of the employment-weighted estimates exhibit small dips around the 1990-91

recession. In the West, employment-weighted earnings also fall from 1992-96; in each of the other

regions, employment-weighted earnings are stagnant after 1992. In contrast, the regression-based

estimates fall more sharply in the early 1990s (in the Northeast and West, hourly wages fall by more

than 10 percent from 1989 to 1990). The regression-based estimates for the Northeast and West

indicate a continued fall in wages after the recession. In the Midwest and South, the regression-based

estimates indicate a recovery in wages in the mid-1990s.

Figure 3 compares a similar set of earnings imputations for women with high school diplomas. As

with Figure 2, the selectivity corrected and uncorrected regression-based estimates are generally quite

close. The only noticeable difference appears in the Midwest where the selectivity corrected estimates

are consistently higher than the uncorrected estimates. The regression-based imputations in Figure 3

are closer to the employment-weighted estimates than in Figure 2. In the Northeast and West, the

regression-based imputations fall over most of the period. In the South, the regression-based

16

imputations change relatively little from 1989-96. In the Midwest, the selectivity-corrected imputations increase over the period while the uncorrected estimates fall slightly.

6. Discussion

This study's regression methodology has a number of useful and advantageous features for developing county-level measures of less-educated women's wage opportunities. First, by utilizing the large volume of information from the SEDF, the model reliably estimates the permanent county-to-county variation in low-skill wages across the U.S. Previous studies have not been able to capture variation for narrow skill classifications at anywhere near this level of geographic detail. Second, the study uses data-driven geographic definitions of local labor markets. This increases our confidence that the areas used are neither too large nor too small. Third, the use of individual-level data allows the model to control for demographic differences across counties and to avoid biases associated with aggregation. Fourth, the use of individual-level data also allows the model to account for selectivity from the employment decision. Thus, the estimation results can be used to predict wages that are representative of the opportunities facing all potential low-skill women workers, not just the select subset that are currently employed. Fifth, by incorporating information from the CPS, the model captures some of the changes over time in labor market opportunities. While the methodology does not capture all of the county-level, time-series variation available from the CPS (as a general set of county and time interactions would), it does allow me to update the imputations for all counties over time.

As the analysis indicates, there are also some shortcomings to the empirical approach. One problem with implications for the wage imputations is that even with the vast amount of data from the

17

SEDF, there are still relatively few earnings observations for particular skill groups in a handful of rural counties. On a population basis, this might not be a critical issue because the affected counties account for a minuscule share of the total number of people from any given skill group. In subsequent work I will consider ways that information for areas with few residents can be combined.

Another problem is the instability of the coefficients on the local labor market variables across regions, particularly for women who have not finished high school. In general, the empirical results indicate that low-skill women's earnings are sensitive to changes in local employment opportunities. The elasticities for hourly wages with respect to employment changes range from .3 to 1.3 for women without high school diplomas and from .5 to 1.0 for women with diplomas. The corresponding elasticities for earnings range from slightly positive to -2.3 for women who are not high school graduates and from essentially zero to -1.1 for women who are graduates. Thus, it appears necessary to account for changes over time in local employment conditions when imputing wages. The instability of the coefficient estimates, however, suggests that the imputations may not provide reliable indications of the time-series variation of wages within certain counties

Endnotes

1. If a survey question in the SEDF or CPS was unanswered, the Census Bureau "allocated" a response using a hot deck procedure. Instead of using the allocated information, my study treats these data as missing and drops the corresponding observations. See Lillard et al. (1986) for a thorough discussion of allocation procedures and their potential effects on empirical labor analyses.

2. Two exceptions are the public use Census data assembled by Tolbert and Killian (1987) which use commuting data to group counties into labor market areas and the personal income measures assembled by the U. S. Bureau of Economic Analysis (1998) which account for inter-county commuting patterns in allocating earnings to counties. Hanushek (1981) also carefully considered the proper geographic dimensions of local labor markets but was forced to examine large and somewhat arbitrary groupings of counties because data for smaller areas were unavailable.

3. For confidentiality purposes, the REIS suppresses some information for counties with small numbers of employers. In some instances, the REIS indicates that $50,000 or less was earned in a particular industry for a given year; in these cases, the analysis imputes an earnings figure of $25,000. In all other instances of suppression, the analysis applies the state-level percentage of employment or earnings for the industry in that year to the reported total level of county employment or earnings. The imputed entries are then rescaled so that all of the industry-specific data for the county sums to the county's aggregate employment or earnings.

4. The probit results are interesting in their own right for describing local employment behavior. In subsequent work I will examine county-level employment predictions based on these results.

Table 1. Percentages of employees working in county of residence and in different industries by gender and education, 1989

	Women				Men			
Working in	less than high school	high school or GED	some college	college graduate	less than high school	high school or GED	some college	college graduate
County of residence	85.4	81.1	79.6	75.8	78.0	72.3	72.2	70.3
Agriculture	1.8	1.2	0.9	0.6	6.6	3.8	2.3	1.5
Mining	0.1	0.2	0.3	0.2	1.2	1.3	0.8	0.8
Construction	1.0	1.5	1.5	0.8	14.0	12.5	8.4	3.6
Manufacturing	21.3	16.5	10.0	7.1	24.3	27.1	21.5	17.1
Trans. and utilities	1.9	4.1	4.6	3.0	6.4	9.2	8.7	4.9
Wholesale trade	2.7	3.5	3.2	2.2	5.0	6.2	6.3	5.0
Retail trade	32.9	23.0	17.0	7.2	21.8	15.9	16.8	7.9
Fin., ins., real est.	3.3	9.9	11.4	7.8	1.6	2.4	5.3	10.0
Services	27.2	27.2	35.8	38.9	12.8	11.6	16.0	28.4
Federal government	1.2	2.9	3.8	3.2	1.0	2.9	4.5	4.9
St./loc. government	6.7	10.0	11.5	29.1	5.3	7.1	9.3	16.0

Note: Data for non-institutionalized, civilian workers aged 16-64 from SEDF. Calculations use sample weights from SEDF. (DCR source: skillgeodist.lst).

Table 2. Trends in earnings and employment

	1989	1990	1991	1992	1993	1994	1995	1996
Employment per capita								
Total	0.545	0.547	0.540	0.536	0.541	0.550	0.560	0.566
Manufacturing	0.081	0.079	0.075	0.073	0.073	0.073	0.073	0.073
Retail trade	0.092	0.092	0.090	0.090	0.091	0.094	0.096	0.097
Service	0.151	0.155	0.157	0.159	0.163	0.166	0.171	0.175
Real earnings per worker (in thousands of 1998 dollars)								
Total	29.7	29.6	29.2	30.1	30.1	30.0	29.8	29.9
Manufacturing	41.2	40.9	40.9	42.5	42.5	43.0	43.0	42.9
Retail trade	17.8	17.4	17.1	17.2	17.1	17.0	16.8	16.6
Service	26.0	26.1	25.7	26.6	26.2	26.2	26.2	26.3
Employment-weighted estimates for women who have not completed high school								
Employment	0.033	0.033	0.033	0.032	0.033	0.033	0.034	0.034
Real earnings	26.9	26.7	26.4	27.2	27.0	26.9	26.7	26.8
Employment-weighted estimates for women who have completed high school								
Employment	0.078	0.078	0.077	0.077	0.078	0.079	0.081	0.082
Real earnings	28.4	28.3	28.1	29.2	29.1	29.0	28.9	29.0
Employment-weighted estimates for women who have completed some college								
Employment	0.085	0.086	0.086	0.085	0.087	0.088	0.090	0.091
Real earnings	28.7	28.7	28.5	29.6	29.6	29.4	29.4	29.5
Employment-weighted estimates for women who completed college								
Employment	0.059	0.060	0.060	0.059	0.060	0.061	0.062	0.063
Real earnings	29.9	30.0	29.8	30.7	30.6	30.5	30.5	30.6

Note: Estimates are population-weighted averages across counties using employment weights from

the SEDF and total and industry-specific earnings and employment data from the REIS. (DCR source: table2.lst).

Table 3. Region-specific earnings regressions for women who have not completed high school.

Variable	Ln(hourly earnings)				Ln(weekly earnings)			
	Northeast	Midwest	South	West	Northeast	Midwest	South	West
Employment	8.081	37.201***	8.624**	17.169**	-8.663	41.124***	1.926	31.810***
	(8.110)	(11.307)	(3.487)	(6.858)	(9.374)	(13.687)	(4.339)	(7.513)
Ln(earnings)	-1.361***	-2.238***	0.046	0.516	-1.085***	-2.260***	-0.049	0.368
	(0.388)	(0.429)	(0.191)	(0.422)	(0.414)	(0.530)	(0.228)	(0.487)
Employment growth	-26.022	-37.528**	-13.929**	-4.388	19.913	-45.644**	-7.500	10.332
	(20.713)	(17.234)	(5.563)	(12.430)	(23.304)	(21.461)	(6.870)	(18.515)
Age	0.123***	0.105***	0.079***	0.092***	0.104***	0.142***	0.080***	0.109***
	(0.023)	(0.015)	(0.007)	(0.035)	(0.024)	(0.018)	(0.009)	(0.038)
Age squared (/100)	-0.156***	-0.132***	-0.101***	-0.114**	-0.126***	-0.179***	-0.098***	-0.135**
	(0.032)	(0.021)	(0.011)	(0.048)	(0.033)	(0.027)	(0.014)	(0.053)
African origin	-0.084***	-0.349***	-0.145***	-0.044	0.032	-0.375***	-0.174***	-0.103
	(0.032)	(0.089)	(0.013)	(0.120)	(0.034)	(0.114)	(0.016)	(0.132)
Hispanic origin	-0.353***	-0.186***	-0.203***	-0.135***	-0.205***	-0.212***	-0.197***	-0.129***
	(0.072)	(0.055)	(0.022)	(0.022)	(0.076)	(0.063)	(0.025)	(0.025)
Other non-white origin	-0.285***	-0.150**	-0.054***	-0.091*	-0.144***	-0.137*	-0.011	-0.056
	(0.046)	(0.065)	(0.015)	(0.050)	(0.047)	(0.074)	(0.019)	(0.055)
8	1.685***	1.260***	0.887***	0.914	0.927**	1.407***	0.433***	0.858
	(0.437)	(0.282)	(0.142)	(0.594)	(0.458)	(0.361)	(0.178)	(0.658)
R^2	0.100	0.107	0.075	0.088	0.097	0.109	0.089	0.091

Note: Data from combined SEDF and CPS files with 57,404 observations from Northeast, 105,379 observations from the Midwest, 190,081 observations from the South, and 41,837 observations from the West. All models also include time and county variables and use sampling weights. White-consistent standard errors in parentheses. (DCR sources: CNC_cprob_r#f1.lst).

*** Significant at .01 level.   ** Significant at .05 level.   * Significant at .10 level.

Table 4. Region-specific earnings regressions for women who have completed high school

Variable	Ln(hourly earnings)				Ln(weekly earnings)			
	Northeast	Midwest	South	West	Northeast	Midwest	South	West
Employment	9.028***	5.572***	7.634***	12.445***	3.974*	6.023***	8.302***	10.550***
	(1.753)	(1.304)	(1.232)	(3.584)	(2.232)	(1.688)	(1.463)	(3.909)
Ln(earnings)	-1.708***	-0.187	-0.501***	-0.227	-1.117***	-0.159	-0.378**	-0.018
	(0.207)	(0.146)	(0.131)	(0.190)	(0.254)	(0.187)	(0.158)	(0.230)
Employment growth	-27.349***	4.041	0.205	-4.643	-18.236***	4.454	0.264	1.689
	(3.902)	(2.579)	(1.966)	(3.494)	(4.997)	(3.344)	(2.354)	(5.118)
Age	0.071***	0.083***	0.093***	0.080***	0.053***	0.096***	0.100***	0.082***
	(0.005)	(0.004)	(0.003)	(0.011)	(0.007)	(0.005)	(0.004)	(0.013)
Age squared (/100)	-0.094***	-0.112***	-0.125***	-0.102***	-0.064***	-0.132***	-0.132***	-0.103***
	(0.009)	(0.007)	(0.006)	(0.018)	(0.011)	(0.009)	(0.007)	(0.021)
African origin	-0.013	-0.225***	-0.179***	-0.231***	0.040***	-0.184***	-0.182***	-0.137*
	(0.011)	(0.029)	(0.007)	(0.066)	(0.013)	(0.035)	(0.009)	(0.073)
Hispanic origin	-0.196***	-0.144***	-0.128***	-0.072***	-0.085***	-0.139***	-0.116***	-0.069***
	(0.028)	(0.030)	(0.013)	(0.022)	(0.033)	(0.038)	(0.015)	(0.025)
Other non-white origin	-0.245***	-0.228***	-0.156***	-0.114***	-0.078**	-0.186***	-0.132***	-0.065**
	(0.033)	(0.027)	(0.014)	(0.027)	(0.037)	(0.032)	(0.016)	(0.030)
8	1.101***	1.288***	1.190***	0.937***	0.337	1.384***	0.980***	0.686**
	(0.192)	(0.134)	(0.086)	(0.298)	(0.234)	(0.169)	(0.102)	(0.349)
R^2	0.125	0.104	0.120	0.120	0.089	0.075	0.108	0.106

Note: Data from combined SEDF and CPS files with 225,158 observations from Northeast, 437,935 observations from the Midwest, 411,746 observations from the South, and 110,818 observations from the West. All models also include time and county variables and use sampling weights. White-consistent standard errors in parentheses. (DCR sources: CNC_cprob_r#f2.lst).

*** Significant at .01 level.   ** Significant at .05 level.   * Significant at .10 level.

Figure 1. Trends in mean annual earnings

Mean Annual Earnings
Employed men aged 18+

Mean Annual Earnings
Employed women aged 18+
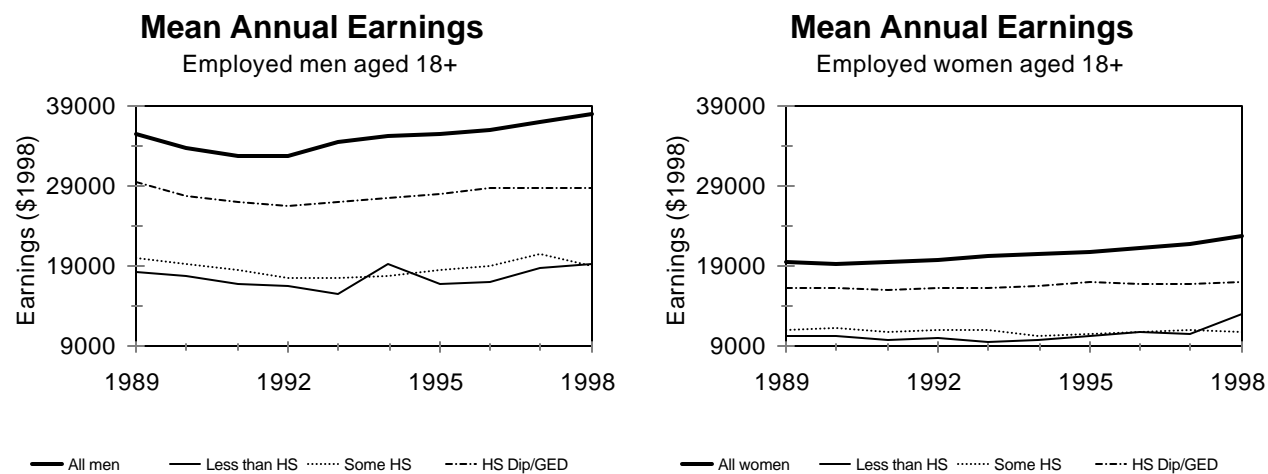
Note: The earnings figures are based on CPS data reported by Census Bureau at
<http://www.census.gov/hhes/income/histinc/incperdet.html> (November 10, 1999). The
schooling definitions before and after 1991 are not entirely consistent.

Figure 2. Alternative wage imputations for women who have not completed high school
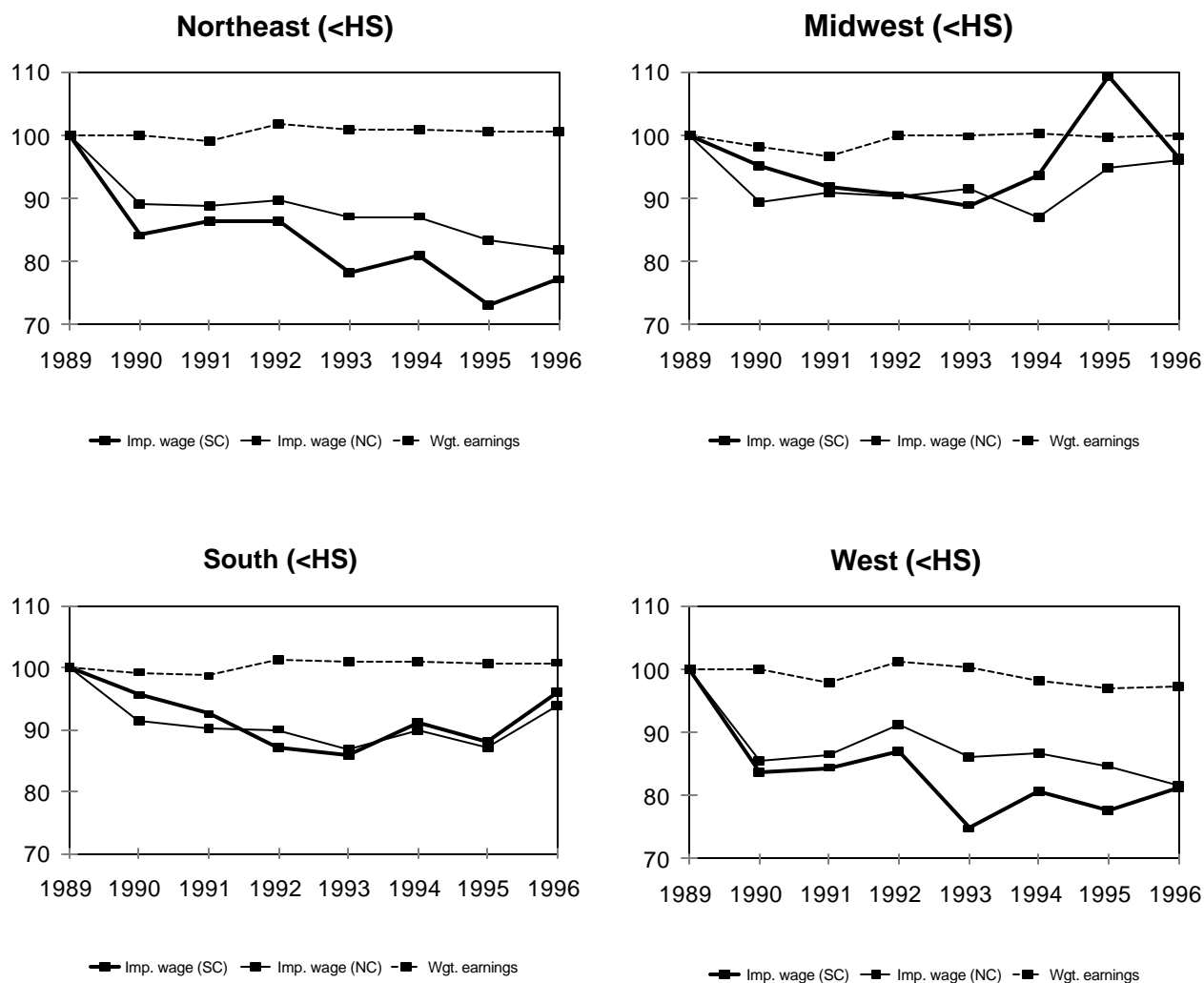
Northeast (<HS)

Imp. wage (SC)　Imp. wage (NC)　Wgt. earnings

Midwest (<HS)

Imp. wage (SC)　Imp. wage (NC)　Wgt. earnings

South (<HS)

Imp. wage (SC)　Imp. wage (NC)　Wgt. earnings

West (<HS)

Imp. wage (SC)　Imp. wage (NC)　Wgt. earnings

Note:　Figures based on population-weighted averages across counties using regression coefficients from Table 3, including unreported county and year coefficients, and employment-weighted data from the REIS. (DCR source: impute_wage_f1#.lst).

Figure 3. Alternative wage imputations for women who have completed high school
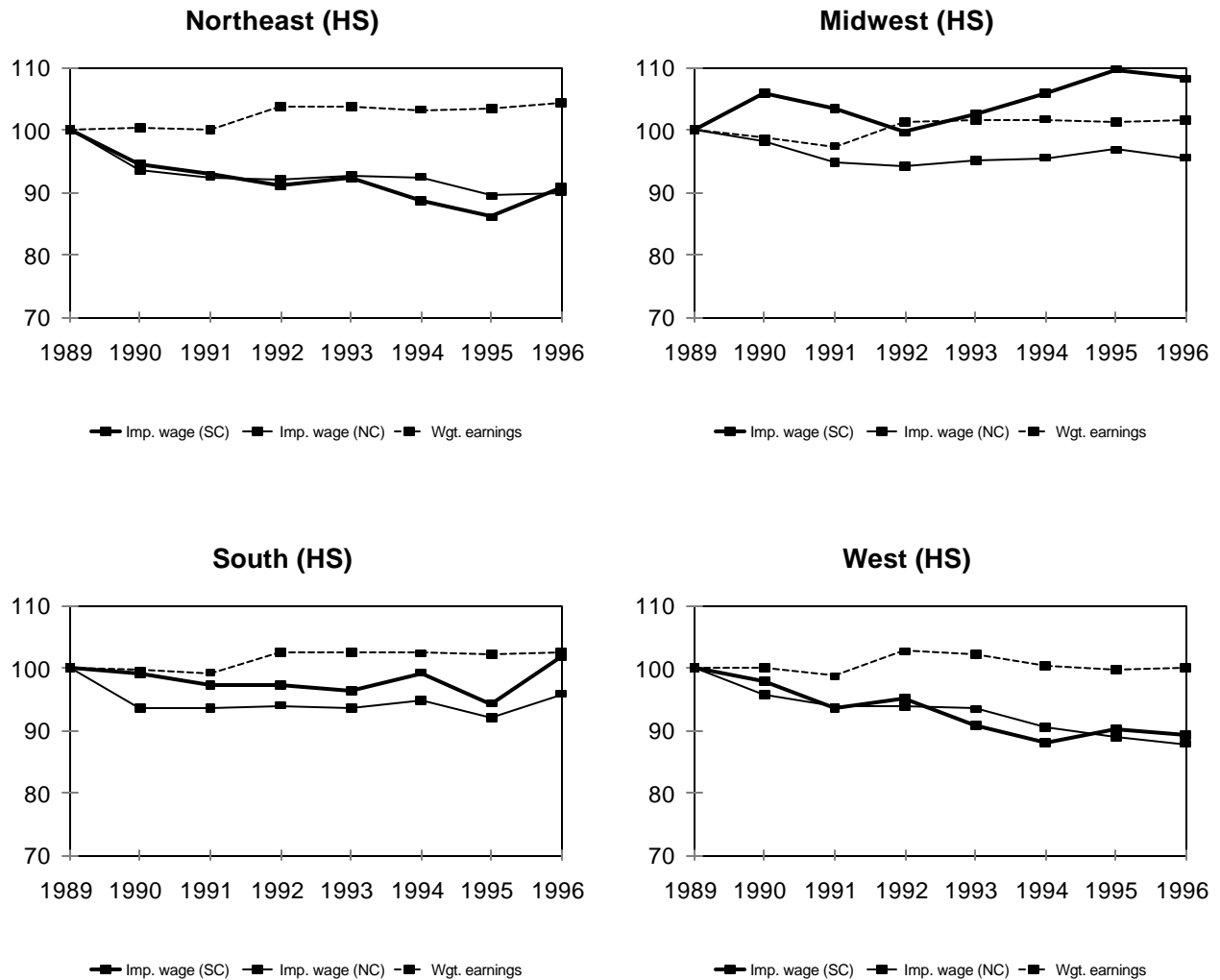
Northeast (HS)

Legend: Imp. wage (SC), Imp. wage (NC), Wgt. earnings

Midwest (HS)

Legend: Imp. wage (SC), Imp. wage (NC), Wgt. earnings

South (HS)

Legend: Imp. wage (SC), Imp. wage (NC), Wgt. earnings

West (HS)

Legend: Imp. wage (SC), Imp. wage (NC), Wgt. earnings

Note: Figures based on population-weighted averages across counties using regression coefficients
from Table 4, including unreported county and year coefficients, and employment-weighted
data from the REIS. (DCR source: impute_wage_f2#.lst).

27

References

Bartik, Timothy J. "The Effects of Local Labor Demand on Individual Labor Market Outcomes for Different Demographic Groups and the Poor," Kalamazoo, MI: W.E. Upjohn Institute for Employment Research, September 1993.

Bartik, Timothy J. "Aggregate Effects in Local Labor Markets of Supply and Demand Shocks," Kalamazoo, MI: W.E. Upjohn Institute for Employment Research, July 1999.

Bartik, Timothy J., and Randall W. Eberts. "Examining the Effect of Industry Trends and Structure on Welfare Caseloads," mimeo, Kalamazoo, MI: W.E. Upjohn Institute for Employment Research, February 1999.

Bound, John, and Harry J. Holzer. "Industrial Shifts, Skills Levels, and the Labor Market for White and Black Males," *Review of Economics and Statistics* 75:3 (August 1993), 387-396.

Bowen, William G., and T. Aldrich Finegan. *The Economics of Labor Force Participation*. Princeton, NJ: Princeton University Press, 1969.

Council of Economic Advisors. "Explaining the Decline in Welfare Receipt, 1993-1996: Technical Report," mimeo, Washington, DC: Executive Office of the President of the United States, 1997.

Council of Economic Advisors. "The Effects of Welfare Policy and the Economic Expansion on Welfare Caseloads: An Update," mimeo, Washington, DC: Executive Office of the President of the United States, August 1999.

Figlio, David N., and James P. Ziliak. "Welfare Reform, the Business Cycle, and the Decline in AFDC Caseloads" in *Welfare Reform and the Economy: What Will Happen When a Recession Comes?* edited by Sheldon Danziger (Kalamazoo, MI: Upjohn Institute for Employment Research, 1999).

Fitzgerald, John M. "Local Labor Markets and Local Area Effects on Welfare Duration," *Journal of Policy Analysis and Management* 14:1 (Winter 1995), 43-67.

Freeman, Richard B. "Economic Determinants of Geographic and Individual Variation in the Labor Market Position of Young Persons," in *The Youth Labor Market Problem*, Richard B. Freeman and David A. Wise (eds.), Chicago: The University of Chicago Press, 1982.

Freeman, Richard B. "Employment and Earnings of Disadvantaged Young Men in a Labor Shortage

Economy," in *The Urban Underclass* Christopher Jencks and Paul E. Peterson (eds.), Washington, DC: The Brookings Institution, 1991.

Freeman, Richard B., and William M. Rodgers III. "Area Economic Conditions and the Labor Market Outcomes of Young Men in the 1990s Expansion," NBER working paper no. 7037, April 1999.

Hanushek, Eric A. "Alternative Models of Earnings Determination and Labor Market Structures," *Journal of Human Resources* 16:2 (Spring 1981), 238-59.

Heckman, James J. "Sample Selection Bias as a Specification Error," *Econometrica* 47 (1979), 153-62.

Hoynes, Hilary. "The Employment, Earnings and Income of Less Skilled Workers over the Business Cycle," NBER working paper no. 7188, June 1999.

Juhn, Chinhui, Kevin Murphy and Brooks Pierce. "Wage Inequality and the Rise in the Returns to Skill," *Journal of Political Economy* 101:3 (June 1993), 410-442.

Katz, Lawrence, and Kevin Murphy. "Changes in Relative Wages, 1963-1987: Supply and Demand Factors," *Quarterly Journal of Economics* 107:1 (February 1992), 35-78.

Lillard, Lee; James P. Smith, and Finis Welch. "What Do We Really Know about Wages? The Importance of Nonreporting and Census Imputation," *Journal of Political Economy* 94:3 part 1 (June 1986), 489-506.

Tolbert, II, Charles M. and Molly Sizer Killian. *Labor Market Areas for the United States*. Washington, DC: U.S. Department of Agriculture, August 1987.

Topel, Robert H. "Local Labor Markets," *Journal of Political Economy* 94:3 part 2 (1986), S111-43.

U.S. Bureau of Economic Analysis. *Regional Economic Information System: 1969-1996*. Washington, DC: U.S. Department of Commerce, 1998.

Walters, Stephen J. K. "Business Climate and Measured Poverty: The Evidence Across States," *Atlantic Economic Journal* 18:1 (March 1990), 20-26.

Appendix A1. Means and standard deviations of analysis variables

Variable	Did not complete high school		Completed high school	
	Mean	(Std. dev.)	Mean	(Std. dev.)
Hourly earnings	9.168	(10.260)	10.759	(9.383)
Weekly earnings	316.849	(345.342)	383.180	(319.475)
Weighted employment per capita in county of residence	0.036	(0.011)	0.083	(0.018)
Weighted annual earnings per worker (/1000) in county of res.	25.981	(5.446)	27.488	(5.356)
Weighted employment growth in county of residence	0.0005	(0.001)	0.001	(0.001)
Age	39.454	(13.295)	37.416	(12.103)
African origin	0.161	(0.368)	0.096	(0.295)
Hispanic origin	0.153	(0.361)	0.048	(0.214)
Other non-white origin	0.106	(0.309)	0.039	(0.193)
Year = 1990 (1991 CPS obs.)	0.009	(0.094)	0.011	(0.103)
Year = 1991 (1992 CPS obs.)	0.008	(0.091)	0.009	(0.097)
Year = 1992 (1993 CPS obs.)	0.008	(0.087)	0.009	(0.095)
Year = 1993 (1994 CPS obs.)	0.007	(0.081)	0.008	(0.090)
Year = 1994 (1995 CPS obs.)	0.007	(0.082)	0.008	(0.089)
Year = 1995 (1996 CPS obs.)	0.006	(0.076)	0.007	(0.082)
Year = 1996 (1997 CPS obs.)	0.006	(0.076)	0.007	(0.083)
Observations	394,701		1,185,657	

Note: Data from combined SEDF and CPS files. Estimates use files' sampling weights. (DCR

sources: CNC_creg_f#.lst).

Appendix Table A2. Probit results for employment in previous year for women who have not completed high school.

	Northeast	Midwest	South	West
Employment	8.282	39.862***	12.859**	-1.165
	(10.436)	(12.200)	(5.141)	(6.973)
Ln(earnings)	-0.711*	-1.702***	0.486*	0.815**
	(0.405)	(0.484)	(0.268)	(0.368)
Employment growth	-47.237**	-37.920*	-21.800***	-15.844
	(24.010)	(22.706)	(7.829)	(18.628)
Age	0.082***	0.078***	0.078***	0.093***
	(0.002)	(0.001)	(0.001)	(0.002)
Age squared (/100)	-0.112***	-0.114***	-0.115***	-0.128***
	(0.002)	(0.002)	(0.001)	(0.003)
African origin	-0.090***	-0.415***	-0.120***	-0.188***
	(0.011)	(0.009)	(0.005)	(0.022)
Hispanic origin	-0.244***	-0.152***	-0.144***	-0.017
	(0.012)	(0.014)	(0.008)	(0.011)
Other non-white origin	-0.140***	-0.218***	-0.048***	-0.126***
	(0.013)	(0.013)	(0.008)	(0.010)
Year = 1990	-0.076**	0.026	0.071***	-0.034
	(0.034)	(0.037)	(0.025)	(0.031)
Year = 1991	-0.069*	-0.084**	0.034	-0.035
	(0.041)	(0.042)	(0.026)	(0.035)
Year = 1992	-0.026	-0.033	-0.077***	-0.117***
	(0.035)	(0.039)	(0.027)	(0.038)
Year = 1993	-0.084**	-0.074*	-0.036	-0.254***
	(0.036)	(0.039)	(0.027)	(0.037)
Year = 1994	-0.056	0.046	0.005	-0.116***
	(0.036)	(0.044)	(0.028)	(0.037)
Year = 1995	-0.104**	0.085	-0.016	-0.133***
	(0.040)	(0.053)	(0.031)	(0.039)
Year = 1996	-0.051	-0.115**	-0.007	0.001
	(0.040)	(0.058)	(0.032)	(0.038)
Log likelihood	-84103	-149922	-292108	-58391
Observations	130168	235919	461173	89247

Note: Data from combined SEDF and CPS files. All models also include county dummy variables.

Standard errors appear in parentheses. (DCR sources: CNC_cprob_r#f1.lst).
*** Significant at .01 level.   ** Significant at .05 level.   * Significant at .10 level.

Appendix Table A3. Probit results for employment in previous year for women who have completed high school.

	Northeast	Midwest	South	West
Employment	10.009***	10.420***	11.864***	24.414***
	(2.604)	(2.564)	(2.154)	(3.799)
Ln(earnings)	-1.577***	-0.257	-0.219	-0.381
	(0.222)	(0.291)	(0.242)	(0.288)
Employment growth	-23.783***	6.314	0.125	1.230
	(6.241)	(5.011)	(3.842)	(8.446)
Age	0.049***	0.054***	0.068***	0.062***
	(0.001)	(0.001)	(0.001)	(0.002)
Age squared (/100)	-0.083***	-0.093***	-0.113	-0.102***
	(0.001)	(0.001)	(0.001)	(0.002)
African origin	0.021**	-0.327***	-0.110***	-0.323***
	(0.009)	(0.007)	(0.005)	(0.015)
Hispanic origin	-0.195***	-0.172***	-0.107***	-0.022**
	(0.013)	(0.015)	(0.008)	(0.011)
Other non-white origin	-0.223***	-0.259***	-0.152***	-0.105***
	(0.014)	(0.014)	(0.009)	(0.010)
Year = 1990	0.004	0.118***	0.116***	0.036
	(0.020)	(0.021)	(0.031)	(0.024)
Year = 1991	-0.030	0.142***	0.104**	-0.008
	(0.026)	(0.025)	(0.048)	(0.026)
Year = 1992	0.046**	0.094***	0.119*	0.071**
	(0.022)	(0.023)	(0.068)	(0.030)
Year = 1993	0.061***	0.101***	0.105	-0.014
	(0.022)	(0.025)	(0.082)	(0.029)
Year = 1994	-0.007	0.124***	0.127	-0.045
	(0.022)	(0.027)	(0.094)	(0.028)
Year = 1995	-0.008	0.139***	0.084	0.004
	(0.026)	(0.032)	(0.107)	(0.030)
Year = 1996	0.063**	0.132***	0.149	-0.008
	(0.026)	(0.035)	(0.122)	(0.030)
Log likelihood	-203709	-375075	-372828	-102708
Observations	337824	639681	632666	169264

Note: Data from combined SEDF and CPS files. All models also include county dummy variables.

Standard errors appear in parentheses. (DCR sources: CNC_cprob_r#f2.lst).
^{***} Significant at .01 level.   ^{**} Significant at .05 level.   [*] Significant at .10 level.